**CAPITULO I**

# 1. CONCEPTOS ESTADÍSTICOS E INFORMÁTICOS A UTILIZAR

## INTRODUCCIÓN

La Estadística e Informática son herramientas que nos van a ayudar en el Desarrollo del sistema de Información de pacientes con enfermedades cardiológicas. Por esta razón se detalla a continuación todos los conceptos estadísticos e informáticos que se van a utilizar.

## Estadística

Ciencia que trata de la recopilación, organización, presentación, análisis e interpretación de datos numéricos con el fin de realizar una toma de decisiones más efectiva.

* 1. **Estadística Descriptiva**

El propósito de la Estadística Aplicada es el de obtener conclusiones de una población en estudio, examinando solamente una parte de ella denominada muestra.

Este proceso, denominado *Inferencia Estadística*, suele venir precedido de otro, denominado *Estadística Descriptiva,* en el que los datos son ordenados, resumidos y clasificados con objeto de tener una visión más precisa y conjunta de las observaciones, intentando descubrir de esta manera posibles relaciones entre los datos, permitiendo sugerir o aventurar cuestiones a analizar en mayor profundidad, ya que contiene técnicas para organizar los datos en una distribución de frecuencias, representarlos en una gráfica y resumirlos para calcular un promedio o una medida de dispersión.

* + 1. **Niveles de medición** 
       - Nominal
       - Ordinal
       - De Intervalo
       - De Razón

**Nivel Nominal:** A este nivel se le considera el más primitivo, él más bajo, o el más limitado de medición.

*Por ejemplo:*

|  |  |
| --- | --- |
| **Religiones indicadas por la población de EEUU. de 14 años o mayores.** | |
| **Religión** | **Total** |
| Protestante | 78.952.000 |
| Católico | 30.669.000 |
| Judío | 3.868.000 |
| Otra Religión | 1.545.000 |
| Ninguna Religión | 3.195.000 |
| Religión no Indicada | 1.104.000 |
| **Total** | 119.333.000 |

Los términos nivel nominal de medición y escala nominal se emplean para hacer referencia a los datos que sólo pueden clasificarse en categorías. Sin embargo, en el sentido exacto de las palabras, no intervienen mediciones ni escalas. En vez de esto sólo hay cuentas o conteos.

En nivel nominal de medición no existe orden particular para los grupos es decir que en primer lugar se podría haber ubicado la católica, en segundo lugar la judía, etc.

Estas categorías se las considera como:

**Mutuamente Excluyente:** Una persona, objeto o medición se incluye solamente en una categoría.

Según la tabla indica que una persona no puede ser Católica y Judía a la vez.

**Exhaustiva:** Cada individuo, objeto o medición debe aparecer en una categoría.

Según la tabla indica que una persona debe pertenecer a una de las categorías.

**Nivel Ordinal:** Una categoría es mayor que la siguiente, esto es, “superior” es una calificación mayor que “bueno”, y, “bueno” es mayor que “promedio”, y así sucesivamente.

|  |  |
| --- | --- |
| **Calificaciones de Estudiantes, semestre de Invierno** | |
| **Calificaciones** | **No. de Calificaciones** |
| Excelentes | 6 |
| Muy bien | 18 |
| Bien | 15 |
| Suficiente | 7 |
| Deficiente | 0 |

La diferencia principal entre nivel de medición nominal y uno ordinal es la relación “mayor que” entre las categorías de nivel ordinal. Por otra parte, la escala ordinal de medición tiene las mismas características que la escala nominal, es decir, las categorías son mutuamente excluyentes y exhaustivas.

**Nivel de Intervalo:** La escala de medición de Intervalo es el siguiente nivel más alto. Incluye todas las características de la escala ordinal, pero además la distancia entre valores es constante.

**Un ejemplo:** es la temperatura en la escala Fahrenheit.

La escala de medición de intervalo tiene las propiedades de ser mutuamente exclusiva y exhaustiva, Orden, Cero significativa (escalas).

**Nivel de Razón**: El nivel de razón (o cociente) es el nivel de medición “más alto”. Este nivel tiene todas las características del de intervalo: las distancias entre números son de un tamaño conocido y constante: las categorías son mutuamente excluyentes.

Las principales diferencias entre los niveles de intervalo y de razón son:

* 1. Los datos de nivel de razón tiene un punto cero significativo.
  2. La razón o cociente de dos números es significativo.
     1. **Distribuciones de Frecuencias**

1. El objetivo de una distribución de frecuencias es organizar datos no agrupados (originales) en alguna forma significativa.
2. Una distribución de frecuencias es un agrupamiento de datos en clases que muestran el número de valores que contiene cada uno.

*El procedimiento es:*

* 1. Se elabora una ordenación, que es una lista de los valores ordenados de menor a mayor, o viceversa.
  2. Se decide el tamaño del intervalo de clase. Si se ha establecido el número de clases el intervalo de clases sugerido suele determinarse por medio de:



Si no se tiene la seguridad del número de clases que se van a utilizar, el intervalo de clases sugerido puede determinarse utilizando esta fórmula:



* 1. Se marcan los datos originales en las marcas adecuadas para elaborar la distribución de frecuencias.

1. Otros criterios para elaborar una distribución de frecuencias.
   1. Evite tener muy pocas o demasiadas clases.
   2. El ancho de los intervalos de clase debe ser igual, si es posible.
   3. Deben evitarse las clases de extremos abiertos, si fuera posible.

***Ejemplo:*** El Gerente Regional de ventas realizó una encuesta a 50 clientes con la siguiente pregunta: ¿Cuántas veces ha realizado compras en esta empresa durante los últimos 30 días? El número de visitas fue:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Datos Originales** | | | | |  | **Datos Originales Ordenados** | | | | |
| 5 | 3 | 3 | 1 | 4 |  | 1 | 3 | 4 | 6 | 8 |
| 4 | 5 | 6 | 4 | 2 |  | 1 | 3 | 4 | 6 | 8 |
| 6 | 6 | 6 | 7 | 1 |  | 1 | 3 | 5 | 6 | 9 |
| 14 | 1 | 2 | 4 | 4 |  | 1 | 4 | 5 | 6 | 9 |
| 4 | 5 | 6 | 3 | 5 |  | 1 | 4 | 5 | 6 | 10 |
| 3 | 4 | 5 | 6 | 8 |  | 2 | 4 | 5 | 6 | 11 |
| 4 | 7 | 6 | 5 | 9 |  | 2 | 4 | 5 | 6 | 12 |
| 11 | 3 | 12 | 4 | 7 |  | 2 | 4 | 5 | 7 | 12 |
| 6 | 5 | 15 | 1 | 10 |  | 3 | 4 | 5 | 7 | 14 |
| 1 | 8 | 9 | 2 | 12 |  | 3 | 4 | 6 | 7 | 15 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Distribución de Frecuencias** | | | | | |
| **Intervalo**  **de Clases** | **Frec.**  **de clases** | **Pto.**  **Medio** | **Frecuencias Acumuladas** | | **Frec.**  **Relativa** |
| **Menos De** | **Mas De** |
| (1 - 4) | 22 | 2,5 | 22 | 50 | 0,44 |
| (5 - 8) | 20 | 6,5 | 42 | 28 | 0,4 |
| (9 - 12) | 6 | 10,5 | 48 | 8 | 0,12 |
| (13 - 16) | 2 | 14,5 | 50 | 2 | 0,04 |
|  | **50** |  |  |  |  |

* + 1. **Polígonos de Frecuencias Acumuladas**
    2. Un polígono de frecuencias acumuladas *“menos de”* permite determinar cuántas o que porcentaje de las observaciones es menor que cierto valor.

|  |
| --- |
|  |

* + 1. Un polígono de frecuencias acumuladas *“más de”* se elabora acumulando las frecuencias de clase empezando con la más elevada. Se gráfica luego los límites verdaderos inferiores y las frecuencias acumuladas. A partir del polígono es posible determinar cuántos o que porcentaje, de los valores son mayores que una cantidad seleccionada.

|  |
| --- |
|  |

* + 1. **Representación Grafica de una Distribución de Frecuencias**

1. Un *Histograma* representa el número de frecuencias de cada clase en forma de barras.

|  |
| --- |
|  |

1. Un *Polígono de Frecuencias* y un *Polígono de Frecuencias Relativas* tienen las clases colocadas en el eje X y las frecuencias de clase en el eje Y. El punto medio de una clase y su frecuencia correspondiente se ubican en un punto representativo. Los puntos se unen para formar el polígono. El área bajo el polígono, como en el histograma, es igual al número total de frecuencias.

|  |
| --- |
|  |

## Estadística Inferencial

La Estadística Inferencial es un conjunto de métodos empleados para determinar algo acerca de una población, con base en una muestra.

## Medidas de Tendencia Central

**Media Aritmética**

La media es el promedio aritmético de las observaciones.

**Media de una Muestra**

La medida de tendencia central (promedio) de uso más amplio es la llamada media aritmética que, por lo general, se designa sólo como media.

Para datos originales, esto es, no agrupados, la media es la suma de todos los valores dividida entre el número total de valores. Su fórmula es la siguiente:



La notación abreviada del álgebra para la media es:

****

= significa media muestral, y se lee “X con barra”

X = indica un valor específico

n = es el número total de valores en la muestra

= es la letra griega sigma mayúscula e indica la operación de sumar un conjunto de datos (sumatoria).

 = indica la suma de todas las X.

**Media de una Población**

Muchos estudios utilizan todos los valores. La media de la población en términos de símbolos es:



En donde:

= indica la media de la población. Es la letra griega mu minúscula.

N = es el número total de observaciones en la población.

**Propiedades de la Media Aritmética**

* 1. Todo conjunto de datos de nivel de intervalo y de nivel de razón tiene una media
  2. Al evaluar la media se incluyen todos los valores.
  3. Un conjunto de datos sólo tiene una media. Esta es única.
  4. La media es una medida muy útil para comparar dos o más poblaciones.
  5. La media aritmética es la única medida de tendencia central en donde la suma de las desviaciones de cualquier valor con respecto a la media siempre será cero.

Expresado en forma simbólica:



**Desventajas de la Media Aritmética**

1. Si uno o dos de los elementos de la muestra o población es muy grande o muy pequeño, la media podría no ser un promedio adecuado para representar los datos.
2. Es inadecuada si hay una clase de extremo abierto para datos agrupados en una distribución de frecuencias.

**Mediana**

Punto medio de los valores después de ordenados de menor a mayor o de mayor a menor. Hay tantos valores por encima de la mediana como por debajo de ella en la ordenación de datos.

Una forma muy fácil de localizar la posición del elemento medio para datos no agrupados es con la fórmula:

Ubicación de la mediana = 

En donde *n* es el número total de elementos

**Propiedades de la Mediana**

1. La mediana es única; esto es, a semejanza de la media sólo existe una mediana para un conjunto de datos.
2. No es difícil determinarla para datos no agrupados. Tan sólo se necesita ordenarlos de menor a mayor o viceversa, y encontrar el valor del elemento central.
3. No se ve afectada por valores muy grandes o muy pequeños, y por lo tanto, es una medida valiosa de la tendencia central cuando ocurre este tipo de valores.
4. Puede calcularse para una distribución de frecuencias de extremo abierto si la mediana no se encuentra en una clase de tal extremo.
5. Puede calcularse para datos de nivel de razón, de intervalo y ordinal.

**Moda**

Es el valor de observación que ocurre con más frecuencia. Puede determinarse la moda para todos los niveles de datos: nominal, ordinal, de intervalo y de relación. La moda también tiene la ventaja de no verse afectada por valores muy altos o muy bajos. Al igual que la mediana, puede utilizarse como medida de tendencia central para distribuciones de extremo abierto.

Sin embargo, la moda tiene algunas desventajas, que hacen que se utilicen con menos frecuencia que la media o la mediana. Para muchos conjuntos de datos, no existe moda porque ningún valor aparece más de una vez. Por ejemplo, no hay moda para este conjunto de datos: 19, 21, 23, 20 y 18. Puesto que cada valor es diferente, podría argumentarse que cada valor es la media. Por el contrario, para algunos conjuntos de datos hay más de una moda. Supóngase que las edades de un grupo son: 22., 26, 27, 27, 31, 35 y 35. Tanto las edades 27 como 35 son modas. Sería cuestionable utilizar la dos modas para representar la tendencia central de este conjunto de datos de edades.

## Cuartiles

*Primer Cuartil Q1*: mediana de datos que preceden a la mediana

*Segundo Cuartil Q2*: mediana

*Tercer Cuartil Q3*: mediana de datos que siguen a la mediana.

**Percentiles**

El percentil es el porcentaje de una distribución que es igual o menor a un número en particular; por ejemplo, considere el diagrama del crecimiento estándar para niñas desde el nacimiento hasta los 36 meses de edad. Para las niñas de 21 meses, el percentil 95 de peso es de 12 Kg., que indica el 95% de 21 meses de edad pesan 12 Kg. o menos y solo 5% pesa más de 12 Kg.

* + 1. **Medidas de Dispersión y Asimetría**

**Amplitud Total**

La medida de dispersión más sencilla es la amplitud total. Se trata de la diferencia entre los valores mayor y menor de un conjunto de datos. Expresada como ecuación:

Amplitud total = Valor Mayor – Valor Menor

**Desviación Media**

Un defecto importante de la amplitud total es que se basa sólo en dos valores, el mayor y el menor: no toma en consideración todos los datos. La desviación media si lo hace. Denominada también como desviación promedio, mide el promedio en donde los valores de una población, o muestra, varían con respecto a su media. En términos de definición, la Desviación Media es la media aritmética de los valores absolutos de las desviaciones con respecto a la media aritmética.



En donde:

x = es el valor de cada observación

 es la media aritmética de los valores

n = es el número de observaciones en la muestra

 = valor absoluto. Es decir no se toma en cuenta los signos de las desviaciones respecto a la media.

**Varianza y Desviación Estándar**

La varianza y la desviación estándar se basa en las desviaciones con respecto a la media.

***Varianza:*** Media aritmética de las desviaciones cuadráticas con respecto a la media.

***Desviación Estándar:*** Raíz cuadrada de la varianza.

***Varianza Poblacional***

Las fórmulas para la varianza poblacional y para la varianza muestral son un poco diferentes.



 es el símbolo para la varianza de una población.

X = es el valor de la observación en la población.

= es la media de la población.

N = es el número total de observaciones en la población.

Para realizar cálculos donde la población es pequeña vale utilizar la fórmula anterior, pero cuando es grande y la media no es un número entero se debe utilizar la fórmula siguiente la cual no se basa en las desviaciones respecto a la media, sino más bien en los valores reales, eliminando así un gran número de restas.



***Desviación Estándar Poblacional***

 o bien 

***Varianza Muestral***

******

En donde:

S2 = es el símbolo empleado para representar la varianza muestral

X = es el valor de las observaciones en la muestra

  = es la media de la muestra

n = es el número total de observaciones en la muestra

***Desviación Estándar***

La desviación estándar de una muestra se utiliza como un estimador de la desviación estándar de la población.

******

* + 1. **Regla Empírica**

Para la distribución de frecuencias simétricas de campana, aproximadamente 68% de las observaciones se encontrará a más y menos una desviación estándar de la media; aproximadamente 95% de las observaciones se encontrará a más y menos dos desviaciones estándares desde la media; y prácticamente todas las observaciones 99.7% se encontrarán a más y menos tres desviaciones desde la media.

* + 1. **Coeficiente de Variación**

Es una medida útil de la dispersión relativa de los datos y se emplea con frecuencia en ciencias biológicas. El coeficiente de variación (CV) se define como la desviación estándar dividida entre la media y multiplicada por 100%. Produce una medida de variación relativa que es relativa al tamaño de la media. La fórmula es:



* 1. Muestreo

En toda investigación estadística existe un conjunto de elementos sobre los que se toma información. Este conjunto de elementos es lo que se denota con el nombre de población o universo estadístico. Cuando el investigador toma información de todos y cada uno de los elementos de la población estadística se dice que está realizando un ***censo,*** pero si se toma información de una parte de los elementos de la población estadística, este proceso recibe el nombre de ***muestreo***. El conjunto de elementos de los que se toma información en el proceso de muestreo se llama ***muestra*** y el número de elementos que la componen ***tamaño muestral.*** Existen varios tipos de muestreo, dependiendo de que la población estadística sea finita o infinita.

* + 1. **Población, Marco y Muestra**

Inicialmente una población es una colección de elementos acerca de los cuales deseamos hacer alguna inferencia. Esta población inicial que se desea investigar se denomina ***población objetivo,*** pero el muestreo de toda la población objetivo no es siempre posible, ya que se presentan problemas de ausentismo, por lo que la población objetivo se restringe al concepto de ***población investigada***, que es la población que realmente es objeto de estudio.

Una unidad de muestreo puede ser un elemento de la población, pero para seleccionar el conjunto de unidades de muestreo que componen la muestra, será necesario disponer de un listado material de unidades de muestreo. Esta relación de unidades de muestreo, de la que se selecciona la muestra se denomina ***marco.***

En general, una ***muestra*** es una colección de unidades de muestreo seleccionadas de un marco o de varios marcos.

**Objetivos**

Loprimero en toda encuesta es fijar en términos concretos los objetivos de la misma, es necesario establecer los objetivos de la encuesta de una forma clara y concisa, y remitirse a estos objetivos conforma se vaya progresando en el diseño e instrumentación de la encuesta. Al momento de fijar los objetivos es necesario tener presentes determinados factores como son:

* + - * ¿Qué información se necesita para cumplirlos?
      * ¿Cuál es el motivo de la encuesta?
      * Existe información disponible de antemano de encuestas piloto?
      * ¿De qué medios materiales y personales se dispone?
      * Límites presupuestarios y temporales
    1. **Diseño de la Muestra**

Para los propósitos de la selección de la muestra debe ser posible dividir la población en lo que se ha denominado ***unidades de muestreo*** de forma no ambigua. Todo elemento de la población debe pertenecer a una sola unidad de muestreo.

Podríamos citar como puntos importantes del diseño de la muestra los siguientes:

* + - * Especificación de la unidades de muestreo
      * Métodos estadísticos para la depuración del marco
      * Posible utilización de la información complementaria
      * Análisis y determinación del tamaño de la muestra
      * Métodos de selección de la muestra, esto es, tipo de muestreo a utilizar
      * Fórmulas para los estimadores a utilizar
      * Fórmulas para la estimación de errores de muestreo
      * Métodos estadísticos para el tratamiento de la falta de respuesta.
      * Control de otros errores ajenos al muestreo.
    1. **Trabajo de Campo**

Se consideran trabajos de campo aquellos que consisten en la obtención de las medidas de las variables objeto de estudio, asociadas a las unidades de la población sobre las que se realiza la medición. Los elementos que consideramos en la realización de los trabajos de campo son los siguientes:

* + - * Las unidades a medir.
      * Las variables objeto de medida
      * El instrumento de medida
      * La realización de la medida y la instrumentalización necesaria
  1. Análisis Multivariado

En un sentido amplio el análisis multivariante se refiere a todas las técnicas estadísticas que simultáneamente analizan medidas múltiples recogidas sobre cada unidad de análisis (objetos, sujetos, etc.). Una definición más exacta es “conjunto de métodos que analizan las relaciones entre un número razonablemente amplio de medidas (variables), tomadas sobre cada objeto o unidad de análisis, en una o más muestras simultáneamente”. La mayor parte de las técnicas dirigen su atención a las covarianzas, correlaciones u otros índices que reflejan el grado de relación entre tres o más variables. El término objeto o unidad de análisis, para referirnos a las entidades sobre las que se toman las medidas.

* + 1. Análisis Factorial de Correspondencia

Es apto para representar tablas de frecuencia. (Ahora se analizan las variables categóricas). Supongamos que los datos corresponden a 2 criterios de clasificación a los que llamaremos caracteres y poblaciones, los cuales se disponen en una tabla de contingencia o tabla de frecuencias absolutas.

**Caracteres o (Clasificación)**

|  |  |  |  |
| --- | --- | --- | --- |
|  | A1 | A2..... | Ap |
| H1 | f11 | f2…. | F1p |
| H2  **Poblaciones**  .  . | f21 | f2p… |  |
| Hn | fn | fn2… | Fnp |

Uno de los objetivos del Análisis Factorial de Correspondencia (AFC) es obtener una representación geométrica de las poblaciones H1,....Hn es relación con la distribución de frecuencias de los caracteres. La distancia que utiliza AFC es la distancia X2

* 1. **Sistema de Información**

Es un servicio que se ofrece en beneficio de algo, sin importar las herramientas tecnológicas que se utilizan.

Se debe tener en cuenta:

**Infraestructura:** se refiere a la tecnología a utilizar, base de datos, redes, servidores y protocolos.

**Infocultura:** se refiere al tipo de cultura informática que tienen las personas para manejar el Sistema de Información.

**Cadena de Valor:** La cadena de valor nos permite analizar todo el proceso que se tienen en un hospital o clínica, con el fin de determinar los posibles problemas que se van a solucionar con el Sistema de información.

**Sistema de Valores:** El sistema de valores nos proporciona los proveedores que dan la información al Sistema, las entidades que tienen el mismo proceso y los clientes que se benefician del Sistema de Información.

* + 1. **Razones de Sistemas de Información**

Las solicitudes de sistemas de información están motivadas por los siguientes tres objetivos generales:

* + - * Resolver un problema
      * Aprovechar una oportunidad
      * Dar respuestas a directivos
    1. **Análisis de Sistemas de Información**

Es un conjunto de procedimientos o programas relacionados de manera que juntos con un plan lógico forman una sola unidad y realicen un objetivo predefinido en el procesamiento de la Información. Esto se lleva a cabo teniendo en cuenta ciertos principios:

* + - * Debe presentarse y entenderse el dominio de la información de un problema.
      * Defina las funciones que debe realizar el Sistema.
      * Represente el comportamiento del sistema a consecuencias de acontecimientos externos.
      * Divida en forma jerárquica los modelos que representan la información, funciones y comportamiento.

Un Análisis de Sistema se lleva a cabo teniendo en cuenta los siguientes objetivos en mente:

* + - * Identifique las necesidades del Cliente.
      * Evalúe que conceptos tiene el cliente del sistema para establecer su viabilidad.
      * Realice un Análisis Técnico y económico.
      * Asigne funciones al Hardware, Software, personal, base de datos, y otros elementos del Sistema.
      * Establezca las restricciones de presupuestos y planificación temporal.
      * Cree una definición del sistema que forme el fundamento de todo el trabajo de Ingeniería.
    1. **Diseño de Sistemas de Información**

El Diseño de Sistemas es el proceso de aplicar ciertas técnicas y principios con el propósito de definir un proceso o un Sistema, con suficientes detalles para permitir su interpretación y realización física.

La etapa del Diseño del Sistema encierra cuatro etapas:

* El Diseño de los datos
* El Diseño Arquitectónico
* El Diseño de la Interfaz
* El Diseño de procedimientos

Para evaluar la presentación del Sistema se debe tener en cuenta los siguientes diseños:

* Diseño de la Salida
* Diseño de Archivos
* Diseño de Interacciones con la Base de Datos
  1. Sistemas de Información Médica

La salud de calidad es uno de los derechos que debe tener todo ser humano, y es un requisito imprescindible para poder salir de los esquemas de pobreza, según organismos como la ONU y el BIP. En el proyecto a desarrollarse se muestra la importancia de la formación, actualización e información de calidad en el sector salud. Se destaca la importancia de los recursos digitales en el proceso de gestión de la información, la planeación de recursos tanto técnicos como científicos con la finalidad de brindar servicios de salud de calidad.

* + 1. **Objetivos**

Al desarrollar una aplicación informática para la confección de historias clínicas, los resultados que se llegarían a tener serían: un control adecuado de la identificación del paciente evitando duplicidades en los números de las historias clínicas y de control y datos erróneos. Control de la documentación, presencia y normalización de cada documento que configura la historia clínica del paciente. Confección de una base de datos relacional de utilidad científica, clínica, epidemiológica y de gestión. Codificación de signos y síntomas, procedimientos, diagnósticos médicos.

* + 1. **Ventajas**

Llevar un registro en el área cardiológica de los datos personales y clínicos del paciente que pueda complementarse con registros de otras características epidemiológicas.

**Asistenciales:** Los pacientes están perfectamente localizados, ordenados, controladas sus pruebas y sus tratamientos, así como la información clínica disponible en tiempo real.

**Científicas:** El sistema permite crear un registro hospitalario en cardiología exhaustivo que tiene capacidad para control asistencial, agrupar pacientes para estudios epidemiológicos, ensayos clínicos y casos control.

**Sociales:** Los pacientes disponen en tiempo real de cualquier lugar donde se encuentren sus datos con informes exhaustivos y legibles de sus patologías con los que pueda trasmitir a sus médicos de cabecera o especialistas su información clínica.

**Gestión:** Datos estadísticos sobre procedimientos empleados, historias clínicas, ingresos, pacientes, etc. Todos ellos imprescindibles para la elaboración de la memoria anual del Servicio de Cardiología.

* + 1. La Información y cómo afectará al Médico y/o al Paciente

La información no son sólo datos (si acaso, se trata de la transformación de estos datos de una manera que facilite la interpretación), también son emociones, filiación, dependencia y nuevos tipos de relaciones que pueden crearse a partir de la generación de esta información.

La tecnología de la información (Information Technology, IT) es aquella tecnología diseñada para facilitar y mejorar el uso de la información. Sus componentes son:

* + - * Sistemas de Input y Output: Facilitan la interacción con el sistema
      * Software: Instrucciones que permiten realizar determinadas operaciones
      * Instrumentos de comunicación: conectan los sistemas IT
      * Procesador: CPU (Central Processing Unit) y memoria interna
      * Sistemas de almacenamiento

La información es útil en la medida en que es útil para la toma de decisiones. La buena información tiene una serie de características:

* + - * Relevante, precisa y exacta: esto es, pertinente para lo que queremos conseguir.
      * Temporal: antes, no después
      * Complementaria: completa lo que ya sabemos, otorgándole mayor inteligilibilidad (Podemos saber la talla de 50 niños, pero quizás ayude más conocer la media de estatura de esta muestra)

Una de las paradojas de las tecnologías de la información, es que no sólo consiste en instalar una herramienta, por costosa que ésta sea, y esperar a que aparezcan los resultados. La clave del éxito está en saber dónde queremos llegar.

* + 1. **Médicos y acceso a Información de Calidad**

Los profesionales de la salud quieren información pertinente, relevante y fácil de usar y conseguir. Un médico emplea mucho tiempo recogiendo e integrando información clínica, aproximadamente un tercio de su jornada laboral.

Internet puede ayudar no sólo facilitando las búsquedas bibliográficas, sino que implica la organización de la asistencia en general. Esto significa un mayor control en el manejo de historias clínicas, en la prescripción y en la gestión de pruebas complementarias.

Tres ventajas básicas de la automatización de archivos médicos:

* + - * Mejora la logística y la organización de los profesionales
      * Limita errores y costes de manejo
      * Permite realizar análisis sistemáticos de la práctica médica, establecer prioridades y planificar recursos

Desde una perspectiva profesional, el valor de la información depende de lo bien que los proveedores cumplan sus compromisos éticos con la mejor relación coste-beneficio.

Un sistema de informática médica debe estar sostenido por un equipo de trabajo y cumplir los siguientes requisitos.

* + - * Permite extraer información de episodios previos
      * La información es fácil de compartir
      * Facilita la creación de un registro centrado en el paciente
      * Facilita la labor de integrar los diagnósticos y tratamientos del paciente
      * Permite mejorar el grado de comunicación y conocimiento con el paciente.
      * Integra documentación de resultados, observaciones, reacciones, y valoraciones
      * Permite basar las decisiones clínicas
    1. **Paciente**

El acceso del paciente a la información en Internet es una realidad. En la actualidad 69 millones utilizan la red para buscar información de salud. Los factores que más influyen en ellos para utilizar Internet son: la enfermedad de un ser querido (40.1%), cambiar sus hábitos dietéticos (26.3%) y haber sido diagnosticados de una enfermedad crónica (22.1%). A la hora de elegir una Web, el consumidor se fija especialmente en la credibilidad del contenido (85.8%) y fácil acceso a la información (64.7%).

* + 1. **La Relación Médico-Paciente y la Telemedicina**

Las Web con información médica y directorios son útiles, sin duda, pero la salud tiene una parte física que no podemos eludir. Las nuevas aplicaciones han de tender a integrar el entorno informativo y el entorno físico, con herramientas que permitan utilizarlas en el entorno adecuado.

Andersen Consulting ha diseñado el Online Medicine Cabinet (OMC). Se trata de una especie de botiquín, conectado a Internet, con herramientas especiales y tomar signos vitales (tensión arterial, colesterol, glucemia). El aparato puede hacer recomendaciones terapéuticas y permitir la búsqueda de información en la red. El médico responsable puede estar informado de los diferentes resultados.

Finalmente, el lifeshirt (lifeshirt.com) es un monitor adaptado en una camiseta, aprobado por la FDA para la monitorización a distancia y diagnóstico de determinadas enfermedades, tales como arritmias y síndrome de apnea del sueño. Permite recoger hasta 40 indicadores (cardíacos y respiratorios), que son almacenados en un PDA y regularmente se envían a la página personal del enfermo.

Se puede concluir que el acceso a la tecnología ha dejado de ser un problema. El punto clave es saber qué y cuánta información es importante.

* 1. **Base de Datos**
     1. **Modelo Entidad – Relación**

El modelo de datos entidad-relación (E-R) está basado en una percepción del mundo real que consta de una colección de objetos básicos, llamados entidades, y de relaciones entre estos objetos. Una entidad es una (cosa) u (objeto) en el mundo real que es distinguible de otros objetos. Las entidades se describen en una base de datos mediante un conjunto de atributos. Una relación es una asociación entre varias entidades.

Además de entidades y relaciones, el modelo E-R representa ciertas ligaduras que los contenidos de la base de datos debe cumplir. Una ligadura importante es la correspondencia de cardinalidades, que expresa el número de entidades con las que otra entidad se puede asociar a través de un conjunto de relaciones.

* + 1. **Modelos Lógicos**

Los modelos lógicos basados en registros se usan para describir datos en los niveles lógicos y de vistas. Se usan para especificar la estructura lógica completa de la Base de Datos como para proporcionar una descripción de alto nivel de la implementación.

* + 1. **Modelo de Datos Físicos**

El modelo de datos físicos se usa para describir datos en un nivel más bajo. En contraste con el modelo de datos lógico, hay pocos modelos de datos físicos en uso.

El modelo físico de datos captura aspectos de la implementación del sistema de Base de datos.

* + 1. **ADO**

ActiveX Data Objects (ADO) es una tecnología ampliable y de fácil uso para agregar acceso a bases de datos a sus páginas Web. Puede utilizar ADO para escribir secuencias de comandos compactas y escalables que conecten con bases de datos compatibles con Open Database Connectivity (ODBC, Conectividad abierta de bases de datos) y orígenes de datos compatibles con OLE DB.

**Características de ADO**

* + - * Objetos creados independientemente.
      * Ayudas de modificación en lote.
      * Soporte para procedimientos almacenados.
      * Diferentes tipos de cursores.
      * Soporte para retornar múltiples recordsets
      * Aplicaciones libres de Hilos para aplicaciones de servidor Web eficientes.

**OBJETOS ADO**

* **Connection**: Mantiene la información de conexión con el proveedor de datos.
* **Command**: Mantiene información acerca de un comando
* **Error**: Mantiene información acerca de algún error ocurrido en el proveedor
* **Parameter**: Es un parámetro para ser usado en un objeto command.
* **Property**: Características de un objeto.
* **Recordset**: Conjunto de registros retornados desde una consulta con un cursor asociado.
  + 1. **Conexión con una Base de Datos**

El primer paso para tener acceso a la información de una base de datos consiste en establecer una conexión con el origen de datos. ADO proporciona el objeto **Connection**, que puede utilizar para establecer y administrar las conexiones entre sus aplicaciones y las bases de datos de ODBC. El objeto **Connection** incorpora diversas propiedades y métodos que puede utilizar para abrir y cerrar conexiones con bases de datos, y para enviar consultas de actualización de la información.

**Creación del archivo de acceso a Datos (dsn)**

Los controladores de bases de datos (programas que pasan información desde su aplicación Web a una base de datos) utilizan un Nombre de origen de datos (DSN) para encontrar e identificar una base de datos ODBC en particular. Normalmente, el DSN contiene información de configuración de la base de datos, seguridad de usuarios y ubicación, y puede tener la forma de una entrada en el registro del sistema de Windows NT o de un archivo de texto.

Con ODBC puede elegir el tipo de DSN que va a crear: Usuario, Sistema o Archivo.

Los DSN de Usuario y de Sistema residen en el registro del sistema de Windows NT. Los DSN de Sistema permiten que todos los usuarios que han iniciado una sesión en un servidor concreto tengan acceso a una base de datos, mientras que los DSN de Usuario limitan la conectividad con la base de datos a los usuarios que dispongan de las credenciales de seguridad apropiadas. Los DSN de Archivo, que tienen la forma de archivos de texto, proporcionan acceso a varios usuarios y son fácilmente transferibles entre un servidor y otro mediante la copia de los archivos DSN.

**Manejo de Conexiones a Bases de Datos**

Uno de los mayores retos del diseño de una aplicación Web sofisticada de base de datos, como una aplicación de entrada de pedidos en línea que atienda a miles de clientes, es la correcta administración de las conexiones con la base de datos. Abrir y mantener las conexiones con las bases de datos, incluso cuando no se transmita información, puede afectar severamente a los recursos del servidor de base de datos y provoca problemas de conectividad.

Las aplicaciones Web de bases de datos bien diseñadas reciclan las conexiones con la base de datos y compensan los retrasos debidos al tráfico de la red. 

* + 1. **SQL SERVER 2000**

Microsoft® SQL Server™ 2000 es el último lanzamiento de los productos de bases de datos de Microsoft, que aprovecha la sólida base establecida por SQL Server 6.5 y SQL 7. Como la mejor base de datos para Windows NT®, SQL Server es el RDBMS ideal para un amplio espectro de clientes corporativos y productores independientes de software (ISV) inmersa en la creación de aplicaciones empresariales. Las necesidades y requisitos del cliente han dado lugar a innovaciones significativas en el producto SQL Server versión 2000, entre las que se incluyen la facilidad de uso, escalabilidad y fiabilidad, y almacenamiento de datos.

**Objetivos de Diseño de SQL Server**

**Liderazgo e innovación**

Las innovaciones permiten a SQL Server 2000 liderar algunas de las categorías de aplicaciones de más rápido crecimiento dentro del sector de las bases de datos. Entre estas categorías se pueden mencionar el comercio electrónico, informática móvil, automatización de sucursales, aplicaciones de líneas de negocio y depósitos de datos.

Entre las importantes áreas de liderazgo e innovación de Microsoft SQL Server 2000 cabe citar:

* Primera base de datos que se amplía desde los portátiles a la empresa mediante el mismo código base y que ofrece una compatibilidad del código del cien por cien.
* Primera base de datos que soporta la configuración automática y la auto-optimización.
* Primera base de datos con un servidor OLAP integrado.
* Primera base de datos con los servicios de transformación de datos (*Data Transformation Services*, DTS) integrados.
* El marco de almacenamiento de datos de Microsoft (*Data Warehousing Framework*) constituye el primer planteamiento de amplia cobertura para la resolución de los problemas que plantea la utilización de metadatos.
* La primera base de datos que ofrece administración multiservidor para un gran número de servidores.
* Una gran variedad de opciones de duplicación de cualquier base de datos.
* La mejor integración con la familia Windows NT Server, Microsoft Office y BackOffice®.
* Acceso universal a los datos (*Universal Data Access*), la estrategia de Microsoft para permitir el acceso de alto rendimiento a una gran cantidad de fuentes de información.

**Facilidad de uso**

La estrategia de Microsoft estriba en convertir a SQL Server en la base de datos que permita llevar a cabo la creación, administración y distribución de las aplicaciones empresariales de la forma más sencilla. Esto significa proporcionar a los desarrolladores un modelo de programación simple y rápido, eliminar la necesidad de administrar la base de datos en las operaciones habituales y proporcionar herramientas sofisticadas para acometer las operaciones más complejas.

SQL Server 2000 reduce el coste total de propiedad mediante opciones tales como la administración de varios servidores con una única consola; ejecución de trabajos basados en eventos y generación de alertas; seguridad integrada y procedimientos de comandos para realizar tareas administrativas.

**Ampliable y fiable**

Microsoft ofrece un único motor de base de datos ampliable desde un equipo portátil que ejecuta el sistema operativo Windows® 95 o Windows 98, hasta clusters multiprocesador simétricos de varios terabytes de información y que ejecutan Windows 2000 Server Enterprise Edition. Todos estos sistemas mantienen la seguridad y fiabilidad que exigen los sistemas empresariales críticos.

SQL Server es también la elección ideal para los sistemas de almacenamiento de datos y OLTP de la gama alta, ya que dispone de funcionalidades de escalabilidad como bloqueo dinámico a nivel de filas, paralelismo entre consultas, consulta distribuida y mejoras en bases de datos de gran tamaño

**Almacenes de datos**

Microsoft ha establecido un planteamiento de amplia cobertura para el *proceso* completo del almacenamiento de datos. El objetivo es facilitar aún más la creación y el diseño de soluciones económicas de almacenamiento de datos mediante la combinación de tecnologías, servicios y alianzas entre fabricantes.

Entre las innovaciones que se incluyen en SQL Server 2000 destinadas a mejorar el proceso de almacenamiento de grandes cantidades de datos, se encuentran:

* "Plato", un componente primordial para las soluciones empresariales que requieran proceso analítico en línea (*Online Analytical Processing*, OLAP), desde la generación de informes y análisis corporativos hasta el modelado de datos y el soporte en la toma de decisiones.
* *Data Transformation Services* (Servicios de transformación de datos) para importar, exportar y transformar datos.
* Mejoras en el tratamiento de las consultas complejas y bases de datos de gran tamaño (VLDB).
* *Microsoft Repository* (Depósito de Microsoft), una infraestructura común para compartir la información.
* Herramientas visuales de diseño para crear y mantener los diagramas de bases de datos.
* Duplicación integrada, que incluye la actualización en varias instalaciones, para mantener almacenes de datos dependientes.
* Integración de soluciones de terceros.
  1. **Lenguajes de Programación**

Un programa de computadora no es más que un conjunto de instrucciones que sigue una computadora para realizar una tarea específica. Un lenguaje de programación se usa para traducir instrucciones tal y como las entendemos los humanos, a los pasos que la computadora puede entender y seguir.

Cuando se baja al nivel más elemental de una computadora, el microprocesador que está en el núcleo de la funcionalidad de la computadora no entiende algo que no sean instrucciones numéricas. Estas órdenes que el procesador comprende recibe el nombre de lenguaje de máquina.

# Arquitectura de Servicios Web

La siguiente figura representa el funcionamiento general de los servicios Web:

|  |
| --- |
| **Figura 1.1**  **Arquitectura general servicios Web** |
| arquitectura-gral |

Podemos deducir que un servicio Web se registra en un repositorio de servicios, el cliente busca en el repositorio el servicio que necesita y luego lo invoca. De manera más detallada "La arquitectura de los servicios Web es una meta-arquitectura que permite que ciertos servicios de red sean dinámicamente descritos, publicados, descubiertos e invocados en un ambiente de cómputo distribuido.

Los servicios Web son aplicaciones auto-contenidas y modulares que pueden ser:

* Descritas mediante un lenguaje de descripción de servicio, como el lenguaje WSDL (Web Service Description Language)
* Publicadas al someter las descripciones y políticas de uso en algún Registro bien conocido, utilizando el método de registro UDDI (Universal Description, Discovery and Integration).
* Encontradas al enviar peticiones al Registro y recibir detalles de ligamiento (binding) del servicio que se ajusta a los parámetros de la búsqueda.
* Asociadas al utilizar la información contenida en la descripción del servicio para crear una instancia de servicio disponible o proxy.
* Invocadas sobre la red al utilizar la información contenida en los detalles de ligamento de la descripción del servicio.
* Compuestas con otros servicios para integrar servicios y aplicaciones nuevas.

|  |
| --- |
| **Figura 1.2**  **Arquitectura detallada de los servicios Web** |
| arquitectura-detallada |

El servicio Web es construido y luego descrito por medio de WSDL9 y registrado bajo el estándar UDDI10, el cliente busca en el registro UDDI (como si fuese un motor de búsqueda al estilo de Google) y obtiene el descriptor WSDL del servicio que necesita, lo invoca haciendo uso de SOAP11 el cual también es utilizado para comunicar la petición entre los diferentes componentes del servidor que aloja el servicio Web, para entregar una respuesta utilizando nuevamente SOAP.

## Protocolos y Estándares Utilizados en los Servicios Web

**XML eXtensible Markup Language.** Este es un lenguaje de marcado al estilo de HTML que pretende dar las pautas generales para la estructuración de información.

**SOAP (Simple Object Access Protocol).** Es un dialecto de XML el cual permite a las aplicaciones invocar métodos de objetos remotos, así como recibir las respuestas de los mismos.

**WSDL (Web Service Description Language).** Es al igual que SOAP, un dialecto de XML que contiene información acerca de la interfaz, semántica y administración de una llamada a un servicio Web.

**UDDI (Universal, Description, Discovery, and Integration) .** Es un protocolo para describir los componentes disponibles de servicios Web. Este estándar permite a las empresas registrarse en un tipo de directorio sección amarilla de Internet que les ayuda anunciar sus servicios, de tal forma que las compañías se puedan encontrarse unas a otras y realizar transacciones en la Web.

**Formularios Web de ASP.NET**

El marco de trabajo de la página de formularios Web de ASP.NET es un modelo de programación escalable de Common Language Runtime que puede utilizarse en el servidor para generar páginas Web dinámicamente.

Concebido como una evolución lógica de ASP (ASP.NET proporciona compatibilidad sintáctica con las páginas existentes), el marco de trabajo de formularios Web ASP.NET se ha diseñado específicamente para tratar varias deficiencias clave del modelo anterior. En particular, proporciona:

* Capacidad para crear y utilizar controles de la interfaz de usuario reutilizables que puedan encapsular funcionalidades comunes y, así, reducir la cantidad de código que tiene que escribir el programador de una página.
* Capacidad para que los programadores puedan estructurar limpiamente la lógica de la página de forma ordenada (no revuelta).
* Capacidad para que las herramientas de desarrollo proporcionen un fuerte soporte de diseño WYSIWYG (Lo que ve es lo que se imprime) a las páginas (el código ASP existente es opaco para las herramientas).
  + 1. **Macromedia Flash**

Flash es un editor de gráficos vectoriales parecido a programas de diseño como ilustrator, Corel Draw o Freehand. Sin embargo, Flash ha sido diseñado para poder añadir animaciones. Estos se crean a partir de fórmulas matemáticas , lo que hace que no pierdan calidad al mostrase al tamaño diferente al original. Las animaciones de gráficos vectoriales tienen un tamaño menor a las creadas a partir de imagenes de mapas de bits, por lo que tardan menos en transmitirse. Los ficheros comprimidos que exporta Flash se denominan películas Shockwave. Para reproducir una de ellas sólo hay que crear un documento HTML que llame al fichero Shockwave y después enviar el documento HTML al servidor Web. Cualquier navegador, Iexplorer, Netscape, etc... equipado con el reproductor Shockwave Flash (gratuito en el Web de Macromedia) podrá reproducir las películas de Flash.

##### Macromedia Dreamweaver

Dreamweaver es un editor del tipo WYSIWYG, es decir, un editor visual que permite diseñar las páginas sin utilizar código HTML, aunque Dreamweaver dispone también del editor HTML o permite configurar un editor HTML externo si así lo deseamos. Para ver el código HTML pulsamos F 10 y se abrirá la ventana del editor, en la misma ventana tenemos el botón del editor externo para utilizar otro programa de edición HTML.

Dreamweaver es el poder combinado con la velocidad y herramientas de la producción que un diseño de páginas Web debe tener.

La plataforma en la cual puede trabajar Dreamweaver es Windows 98 SE/Me/NT/2000/XP.

* + 1. **Macromedia Fireworks**

Fireworks es la herramienta de desarrollo usada en la producción de páginas Web compleja en vías de desarrollo diseña y se integra con otros productos de Macromedia como Dreamweaver.

La integración entre Fireworks y Dreamweaver refuerza la productividad de diseñar páginas Web. Firewords MX, un resultado profesional es la integración con otros ambientes, ya que es como un vector que cambie de tamaño y apariencia a los elementos gráficos de forma sencilla, ofreciendo una alta calidad en la perfección de sus gráficos.

* + 1. **Costos de los Productos Macromedia**

|  |  |  |
| --- | --- | --- |
| **Tabla 1.1**  **Costos de los Productos Macromedia** | | |
| **Herramientas de autoría y de medios interactivos** | **Completo (desde)** | **Actualización (desde)** |
| Macromedia Studio MX 2004 | [999 €](http://www.macromedia.com/go/ols_eu_studio) | [479 €](http://www.macromedia.com/go/ols_eu_studio) |
| Macromedia Studio MX 2004 con Flash Pro | [1.129 €](http://www.macromedia.com/go/ols_eu_studio_pro) | [569 €](http://www.macromedia.com/go/ols_eu_studio_pro) |
| Dreamweaver MX 2004 | [479 €](http://www.macromedia.com/go/ols_eu_dreamweaver) | [239 €](http://www.macromedia.com/go/ols_eu_dreamweaver) |
| Macromedia Flash MX 2004 | [599 €](http://www.macromedia.com/go/ols_eu_flash) | [239 €](http://www.macromedia.com/go/ols_eu_flash) |
| Macromedia Flash MX Professional 2004 | [839 €](http://www.macromedia.com/go/ols_eu_flash_pro) | [349 €](http://www.macromedia.com/go/ols_eu_flash_pro) |
| Fireworks MX 2004 | [359 €](http://www.macromedia.com/go/ols_eu_fireworks) | [179 €](http://www.macromedia.com/go/ols_eu_fireworks) |